Identification of genetic markers and QTL for carcass quality traits within the American Simmental Association Carcass Merit Program[1]

Jordan K. Hieber, Rachel L. Endecott,[2] and Jennifer M. Thomson[3]

Department of Animal and Range Sciences, Montana State University, Bozeman, MT 59717

Transl. Anim. Sci. 2018.2:S39–S43
doi: 10.1093/tas/txy032

INTRODUCTION

With the first matings at the Sheek Ranch in Cabool, MO, in the spring of 1997, the American Simmental Association (**ASA**) launched a program that eventually changed the direction and collective futures of all producers and users of Simmental genetics. Then, simply known as the Carcass Merit Project (**CMP**), the top sires of the Simmental and Simbrah breeds were randomly mated to commercial females with the plan to collect difficult-to-get progeny sire group carcass information and sometimes tenderness data. In addition to information about carcass traits, this program contributes well over 1,000 birth, weaning, and yearling weights, and calving ease scores each year and nearly 4,000 shear force records have been collected. Chute scores have been collected on a portion of the cows in the CMP as well as fertility data. Another more recent addition to the CMP has been the inclusion of intake and feed conversion data on a portion of the calves.

These data allow the ASA to amass substantial information for benchmarking so that performance and value expectations can be conveyed to current and potential customers of ASA's members. From providing confidence regarding levels of calving assistance, all the way to predicting end product value for those wishing to be profitable in the feeding business, ASA can reliably estimate the performance of their genetic products due in part to the information collected on SimGenetics sires tested in these "real world" commercial situations.

In the last 25 yr, a shift has occurred in the U.S. beef industry from a commodity-based market to one that is based on quality or value added. This has been facilitated by the ability of cattlemen to identify and select animals with superior carcass merit. In an effort to retain market share and increase overall consumer acceptance, many producers have placed at least some importance on selection for marbling development or carcass quality. This means that a large number of our current cowherds have females that possess merit in the area of marbling or carcass quality. As a result, there is interest in how carcass-based selection impacts the maternal performance of these animals. In addition, there is a desire to continue to improve genetic merit for carcass and product consistency in the market place. Genomic or marker-based selection is a tool to assist this desired breeding objective and is an area where we still have genetic variation that cannot be predicted with the genomic tools currently available. With that, our hypothesis is that we can determine genetic markers and QTL in carcass traits from the 30 yr of CMP data collected by the ASA.

MATERIALS AND METHODS

Data were gathered by the ASA CMP. All records containing calf birth weights, calving

[1]Appreciation is expressed to The Bair Ranch Foundation for funding this research.

[2]Present address: American Simmental Association, Bozeman, MT 59715

[3]Corresponding author: jennifer.thomson@montana.edu

Received March 16, 2018.
Accepted April 14, 2018.

ease scores, weaning weights, back fat, ribeye area, marbling, HCW, and KPH were extracted from the ASA database. Sire EPD, performance, and genotype data were also extracted.

Progeny data were organized into sire families for all traits, and progeny performance phenotypes were constructed. Sires that had either SNP50K or imputed 50 K data were used in the overall analysis. The following workflow that was used was specifically developed by Golden Helix, utilizing best practices in genetic association analyses. Quality control of samples was done through a series of filters (Golden Helix, 2017). First, samples were filtered out and removed if they have a call rate of ≤0.95. Next, markers were filtered out and removed if they had all of the following: a call rate <0.85, had >2 alleles, and if the minor allele frequency (**MAF**) was <0.01. Then, data were pruned for markers in linkage disequilibrium and to inactive markers in nonautosomal chromosomes.

Quality control of SNP data also excluded SNPs with spurious position, low call rates (<95%), out of Hardy–Weinberg equilibrium ($P < 0.01$), or less than 10% MAF. Samples were then filtered to determine relatedness. A heatmap was produced to illustrate patterns of relatedness between individuals. Principal component analysis (**PCA**) was applied to the similarity matrix, where it was found that the first three eigen vectors represented greater than 50% of the stratification on the SNP data. The calculated relatedness between individuals was used as a covariate in the association analysis.

A single-locus mixed linear model (EMMAX; Kang et al., 2010) in SVS software (Golden Helix, version 8.7.2-2017-08-11) was used to perform a regression-based association analysis on genotype data while correcting for cryptic relatedness and pedigree structure with phenotype as the response variable and how much variation individual markers are accounting for as explanatory variable.

Benjamini–Hochberg multiple comparison corrections were used to minimize false-positive associations. Manhattan plots were created to visualize the association analyses. On the Manhattan plots, the genome-wide significance level for the Benjamini–Hochberg correlation with −log10(*P* value) is 5×10^{-8} (Ehret, 2010), and markers that were above the level of significance were used to identify regions of the genome associated with the trait in question. Regions with clusters of significantly associated markers were then labeled as putative QTL and used to identify potential positional candidate genes within each.

RESULTS AND DISCUSSION

The data contain samples from 3,849 individuals, where some individuals appear to be more closely related to each other than other samples. We were able to group individuals by sire, in which 2,745 individuals had known sires, producing 395 sire families. Sire families ranged in size from 1 progeny up to 150 progeny; however, not all of the progeny had reported information for their carcass traits. For

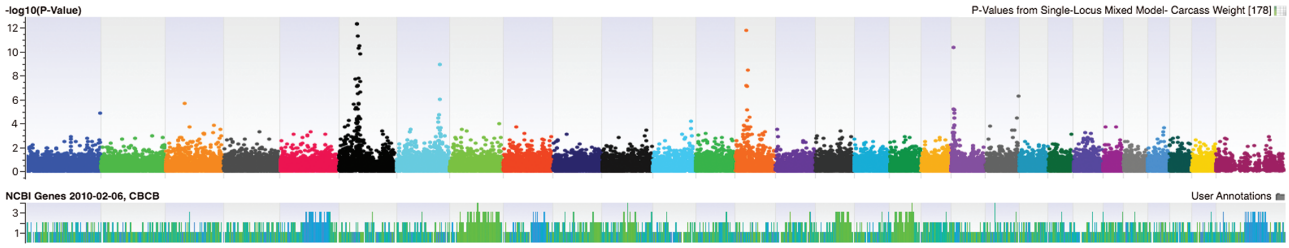

Figure 1. Manhattan plot for HCW. Markers above −log10(*P* value) of 5×10^{-8} are genome-wide association significant markers. Vertical clusters of markers are also of interest as they are indicating suggestive QTL in those regions.

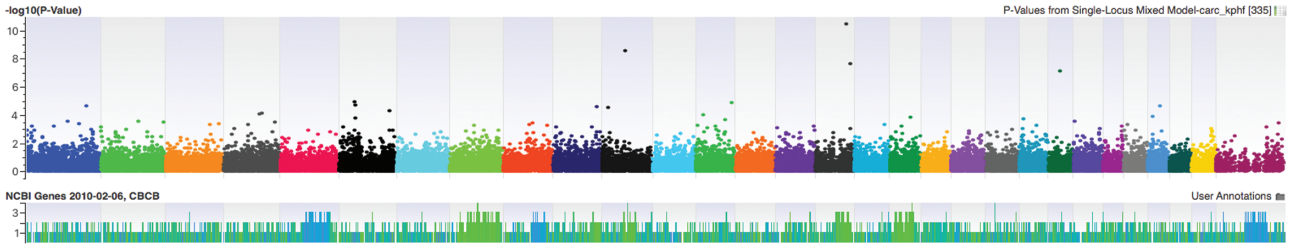

Figure 2. Manhattan plot for carcass KPH. Markers above −log10(*P* value) of 5×10^{-8} are genome-wide association significant markers. Vertical clusters of markers are also of interest as they are indicating suggestive QTL in those regions.

Table 1. Significant HCW genome-wide association markers that are within 100,000 bp

Chromosome	Position (bp)	Positional candidate gene
1	94860836-94882093	
	130032478-130112755	End of *Bos taurus* CLSTN2
3	119048887-119077206	
4	20103064-20181749	Beginning of *Bos taurus* SCIN
6	13918445-13973477	
	17224897-17282916	
	19426158-19451737	*Bos taurus* DKK2
	37294843-37335860	
	37801349-37839427	
	37839427-37925393	*Bos taurus* PPM1K
	39147750-39172862	
	39313672-39346170	
	39503443-39556588	
	39721727-39773228	
	39773228-39837065	
	40819552-40893067	
	41343408-41379490	*Bos taurus* SLIT2
	42239393-42267374	
	42319104-42387759	
	42609559-42628140	
	43303952-43330106	
	44205092-44363683	
	94627787-94998310	
7	31435921-31471496	Beginning of *Bos taurus* CSNK1G3
	80403492-80483871	
	83621039-83648346	*Bos taurus* ACOT12
	86936090-86979103	
	88971675-89026917	
	89496565-89545945	
	90202205-90239425	
	92654719-92719881	
8	8054728-8109244	*Bos taurus* XKR6
12	68227222-68271937	Beginning of *Bos taurus* GPC6
	81445405-81491302	
14	19220744-19290077	
	23853811-23893220	*Bos taurus* SOX17
	25425357-25459674	
	25612510-25698286	
	26542736-26621020	
	26926569-26949215	End of *Bos taurus* TOX
	30329532-30361887	
	50784282-50874869	Beginning of *Bos taurus* TRPS1
15	4094542-4183554	
	6296367-6362468	End of *Bos taurus* MMP20
16	776322–866294	End of *Bos taurus* CHI3L1
19	8216070-8280552	
20	4618689-4622894	
24	37002274-37076014	
	42737372-42762598	
25	37656205-37693781	*Bos taurus* KPNA7
28	25938159-25986353	End of *Bos taurus* TACR2, Beginning of *Bos taurus* TSPAN15
	26710422-26786093	End of *Bos taurus* LRRC20, Beginning of *Bos taurus* EIF4EBP2

Table 2. Significant KPH genome-wide association markers that are within 100,000 bp

Chromosome	Position (bp)	Positional candidate gene
1	883895-950841	*Bos taurus* ATP5O
	16145053-16196001	
	142401535-142446153	
2	30262141-30307800	
6	3093621-3149732	
	27158687-27183822	
	30782962-30832561	Beginning of *Bos taurus* MBPR1B
	41343408-41443081	*Bos taurus* SLIT2
	42155077-42239393	
8	40775647-40800617	
	51330787-51369892	
	73881694-73907982	
	101044054-101135756	*Bos taurus* PALM2
	101135756-101167884	*Bos taurus* PALM2
9	55740550-55802932	
10	6076725-6096967	
	47114150-47197199	
	92952608-92984267	
11	15919622-15945389	*Bos taurus* LTBP1
13	49963611-50004272	
	62881877-62909025	End of *Bos taurus* BPIFB6, Beginning of *Bos taurus* BPIFB3
14	59112331-59139878	Beginning of *Bos taurus* ANGPT1
16	37479436-37505165	
17	64189856-64225341	
19	37670702-37699961	End of *Bos taurus* NXPH3
22	37615930-37652444	
24	56487933-56564480	*Bos taurus* WDR7
25	40022986-40060928	
	41813524-41849369	*Bos taurus* ELFN1
26	6051502-6092833	
	23048759-23129849	*Bos taurus* SUFU
27	10697257-10764825	
	30025162-30089811	
29	41778946-41854768	End of *Bos taurus* POLR2G, *Bos taurus* TAF6L, *Bos taurus* TMEM179B, *Bos taurus* TMEM223, *Bos taurus* NXF1, *Bos taurus* STX5, *Bos taurus* WDR74
	42620218-42696595	End of *Bos taurus* ATL3, Beginning of *Bos taurus* RTN3
	42985739-43043207	*Bos taurus* MACROD1, Beginning of *Bos taurus* NAA40

those who had reported carcass trait information, progeny performance phenotypes were constructed.

To identify potential candidate genes and pathways related to each carcass trait, Manhattan plots were created for each trait and were used to identify regions of the genome that were of interest for this study. In the Manhattan plots, each new color represents a new chromosome, beginning with chromosome 1 on the left and continuing in ascending order to chromosome 29, and each dot represents a marker.

We found 8 out of 261 total chromosomes to have genome-wide association significant markers. For HCW, chromosomes 6, 7, 14, and 20 had 5, 1, 2, and 1 significant markers, respectively (Figure 1). For KPH, chromosomes 11 and 16 each had one significant marker (Figure 2). Chromosome 20 for average HCW, chromosome 16 for average carcass marbling, and chromosome 17 for average carcass fat each had one significant marker.

Although we found a low number of genome-wide significant markers, the areas of the chromosomes with vertical clusters of markers were of interest to us as they indicated suggestive QTL in those regions. On an individual marker basis, there were markers with significant P values ($P < 0.01$) that fall within 100,000 bp of one another, explaining variation in the trait, which are listed in Tables 1–2.

Based on the significant markers that were within 100,000 bp, those bp regions were cross-referenced with the University of California-Santa Cruz Genome Browser to determine whether any positional candidate genes had been previously identified for beef cattle. Not all significant genomic regions had previously reported positional candidate genes indicating that there are carcass traits that are being impacted by genes not yet reported.

CONCLUSION

This study demonstrates the value of the ASA CMP for identifying and characterizing genomic variants that impact carcass traits in Simmental and Simmental-influenced cattle. This research identified eight chromosomes harboring QTL for various carcass traits and identified some previously unreported positional candidate genes. This can be used to improve the accuracy of breeding value estimations and increase the value of genomic data to Simmental producers.

IMPLICATIONS

While these results are important, there is further research needed before these results are immediately applicable to the ASA and producers. First, choromosome-wide associations will be performed to refine the currently identified QTL. Second, markers and regions that have been identified in this project will be checked against already known carcass trait QTL. Next, the amount of variation the markers and QTL explain will be assessed to improve the accuracy of genomic EPD predictions by including them as correlated traits in a prediction of breeding value. Lastly, genetic and phenotypic correlations to other economically important traits, specifically maternal traits, will be calculated. This will allow for the ASA to advise breeding recommendations to maximize producer profitability by determining the best balance of selective pressure to continue to improve carcass traits while minimizing the negative impacts on other traits.

LITERATURE CITED

Ehret, G. B. 2010. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. Curr. Hypertens. Rep. 12:17–25. doi:10.1007/s11906-009-0086-6

Golden Helix. 2017. Bovine GWAS with mixed linear model tools. SNP & Variation Suite Manual v8.x. Bozeman, MT: Golden Helix, Inc. www.goldenhelix.com.

Kang, H. M., J. H. Sul, S. K. Service, N. A. Zaitlen, S. Y. Kong, N. B. Freimer, C. Sabatti, and E. Eskin. 2010. Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 42:348–354. doi:10.1038/ng.548

Segura, V., B. J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, Q. Long, and M. Nordborg. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44:825–830. doi:10.1038/ng.2314